Industrial Mathematics Institute

2001:07

Two lower estimates in greedy approximation

V.N. Temlyakov

IMI

Preprint Series

Department of Mathematics
University of South Carolina

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2001**	2. REPORT TYPE	3. DATES COVERED **00-00-2001 to 00-00-2001**
4. TITLE AND SUBTITLE **Two Lower Estimates in Greedy Approximation**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of South Carolina,Department of Mathematics,Columbia,SC,29208**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a REPORT **unclassified**	b ABSTRACT **unclassified**	c THIS PAGE **unclassified**	**Same as Report (SAR)**	**13**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Two lower estimates in greedy approximation[1]

V.N.Temlyakov

University of South Carolina, Columbia, SC 29208, USA

Abstract. We prove one lower estimate for the rate of convergence of Pure Greedy Algorithm with regard to a general dictionary and another lower estimate for the rate of convergence of Weak Greedy Algorithm with a special weakness sequence $\tau = \{t\}$, $0 < t < 1$, with regard to a general dictionary. The second lower estimate combined with the known upper estimate gives the right (in the sense of order) dependence of the exponent in the rate of convergence on the parameter t when $t \to 0$.

1. Introduction

This paper is a followup to the paper [L] of E. Livshitz. In Sections 2,3 we study the convergence rate of Pure Greedy Algorithm and in Section 4 we study Weak Greedy Algorithm. We define first the Pure Greedy Algorithm (PGA) in Hilbert space H. We describe this algorithm for a general dictionary $\mathcal{D}$. If $f \in H$, we let $g(f) \in \mathcal{D}$ be an element from $\mathcal{D}$ which maximizes $|\langle f, g\rangle|$. We shall assume for simplicity that such a maximizer exists; if not suitable modifications are necessary (see Weak Greedy Algorithm below) in the algorithm that follows. We define

$$G(f,\mathcal{D}) := \langle f, g(f)\rangle g(f)$$

and

$$R(f,\mathcal{D}) := f - G(f,\mathcal{D}).$$

Pure Greedy Algorithm (PGA). *We define $R_0(f,\mathcal{D}) := f$ and $G_0(f,\mathcal{D}) := 0$. Then, for each $m \geq 1$, we inductively define*

$$G_m(f,\mathcal{D}) := G_{m-1}(f,\mathcal{D}) + G(R_{m-1}(f,\mathcal{D}),\mathcal{D})$$

$$R_m(f,\mathcal{D}) := f - G_m(f,\mathcal{D}) = R(R_{m-1}(f,\mathcal{D}),\mathcal{D}).$$

For a general dictionary $\mathcal{D}$ we define the class of functions

$$\mathcal{A}_1^o(\mathcal{D}, M) := \{f \in H : f = \sum_{k\in\Lambda} c_k w_k, \quad w_k \in \mathcal{D},\ \#\Lambda < \infty \text{ and } \quad \sum_{k\in\Lambda} |c_k| \leq M\}$$

and we define $\mathcal{A}_1(\mathcal{D}, M)$ as the closure (in H) of $\mathcal{A}_1^o(\mathcal{D}, M)$. Furthermore, we define $\mathcal{A}_1(\mathcal{D})$ as the union of the classes $\mathcal{A}_1(\mathcal{D}, M)$ over all $M > 0$. For $f \in \mathcal{A}_1(\mathcal{D})$, we define the "semi-norm"

$$|f|_{\mathcal{A}_1(\mathcal{D})}$$

as the smallest M such that $f \in \mathcal{A}_1(\mathcal{D}, M)$.

[1]This research was supported by the National Science Foundation Grant DMS 9970326 and by ONR Grant N00014-91-J1343

It was proved in [DT] that for a general dictionary $\mathcal{D}$ the Pure Greedy Algorithm provides the following estimate

$$\|f - G_m(f,\mathcal{D})\| \le |f|_{\mathcal{A}_1(\mathcal{D})} m^{-1/6}. \tag{1.1}$$

(In this and similar estimates we consider that the inequality holds for all possible choices of $\{G_m\}$.) The paper [DT] contains also an example of a dictionary $\mathcal{D}$ and an element f such that

$$\|f - G_m(f,\mathcal{D})\| > \frac{1}{2}|f|_{\mathcal{A}_1(\mathcal{D})} m^{-1/2}, \quad m \ge 4. \tag{1.2}$$

We proved in [KT] a new estimate

$$\|f - G_m(f,\mathcal{D})\| \le 4|f|_{\mathcal{A}_1(\mathcal{D})} m^{-11/62} \tag{1.3}$$

which improves a little the original one (see (1.1)).

E. Livshitz [L] proved that there exist $\delta > 0$, a dictionary $\mathcal{D}$ and an element $f \in H$, $f \ne 0$, such that

$$\|f - G_m(f,\mathcal{D})\| \ge C m^{-1/2+\delta} |f|_{\mathcal{A}_1(\mathcal{D})}$$

with a positive constant C. We develop and refine ideas from [L] here to prove the following lower estimate.

Theorem 1.1. *There exist a dictionary $\mathcal{D}$ and an element $f \in H$, $f \ne 0$, such that*

$$\|f - G_m(f,\mathcal{D})\| \ge C m^{-1/3} |f|_{\mathcal{A}_1(\mathcal{D})}$$

with a positive constant C.

In Section 4 we study the Weak Greedy Algorithm. Let a sequence $\tau = \{t_k\}_{k=1}^\infty$, $0 \le t_k \le 1$, be given. Following [T] we define Weak Greedy Algorithm as follows.

Weak Greedy Algorithm (WGA). *We define $f_0^\tau := f$. Then for each $m \ge 1$, we inductively define:*

1). $\varphi_m^\tau \in \mathcal{D}$ is any satisfying

$$|\langle f_{m-1}^\tau, \varphi_m^\tau\rangle| \ge t_m \sup_{g\in\mathcal{D}} |\langle f_{m-1}^\tau, g\rangle|;$$

2).

$$f_m^\tau := f_{m-1}^\tau - \langle f_{m-1}^\tau, \varphi_m^\tau\rangle \varphi_m^\tau;$$

3).

$$G_m^\tau(f,\mathcal{D}) := \sum_{j=1}^m \langle f_{j-1}^\tau, \varphi_j^\tau\rangle \varphi_j^\tau.$$

In Section 4 we discuss the following question. How does the weakness sequence τ affect rate of convergence of WGA? We consider here only the special case of weakness sequences $\tau = \{t_k\}_{k=1}^\infty$ with $t_k = t$, $k = 1, 2, \dots$, $0 < t < 1$. In order to stress this we replace in notations τ by t. It is known from [J] that WGA with the above special weakness sequence $\{t\}$ converges for any $0 < t < 1$. We show in Section 4 that the weakness parameter t affects the rate of convergence of WGA on the class $\mathcal{A}_1(\mathcal{D})$. For the WGA we have the following upper estimate [T].

Theorem 1.2. *Let $\mathcal{D}$ be an arbitrary dictionary in H. Assume $\tau := \{t_k\}_{k=1}^\infty$ is a nonincreasing sequence. Then for $f \in \mathcal{A}_1(\mathcal{D}, M)$ we have*

$$\|f - G_m^\tau(f,\mathcal{D})\| \le M(1 + \sum_{k=1}^m t_k^2)^{-t_m/2(2+t_m)} \tag{1.4}$$

for any realization of $G_m^\tau(f,\mathcal{D})$.

In a particular case $\tau = \{t\}$, ($t_k = t$, $k = 1, 2, \dots$), (1.4) gives

$$\|f - G_m^t(f,\mathcal{D})\| \le M(1 + mt^2)^{-t/(4+2t)}, \quad 0 < t \le 1. \tag{1.5}$$

This estimate implies the following inequality

$$\|f - G_m^t(f,\mathcal{D})\| \le C_1(t) m^{-at} |f|_{\mathcal{A}_1(\mathcal{D})}, \quad a < 1/6, \tag{1.6}$$

with the exponent at approaching 0 linearly in t. We prove in Section 4 that this exponent can not decrease to 0 slower than linearly.

Theorem 1.3. *There exists an absolute constant $b > 0$ such that for any $t > 0$ we can find a dictionary $\mathcal{D}_t$ and a function $f_t \in \mathcal{A}_1(\mathcal{D}_t)$ such that for some realization $G_m^t(f_t, \mathcal{D}_t)$ of Weak Greedy Algorithm we have*

$$\liminf_{m\to\infty} \|f_t - G_m^t(f_t,\mathcal{D}_t)\| m^{bt} / |f_t|_{\mathcal{A}_1(\mathcal{D}_t)} > 0. \tag{1.7}$$

2. General formulas

We will be constructing simultaneously two sequences of elements $\{x_n\}_{n=N}^\infty$ and $\{g_n\}_{n=N}^\infty$. A number N will be chosen later to be large enough. Let $\{e_j\}_{j=1}^\infty$ be an orthonormal basis for H. We let $\{x_n\}$ have the following form

$$x_n = \sum_{k=1}^n a_{n,k} e_k, \quad n = N, \dots, \tag{2.1}$$

and the $\{g_n\}$ have the form

$$g_{n+1} = \gamma_{n+1} x_n + h_{n+1} + \xi_{n+1} e_{n+1}, \quad n = N, \dots, \tag{2.2}$$

with the sequences $\{\gamma_k\}$, $\{h_k\}$, $\{\xi_k\}$, $0 \le \xi_k \le 1$, to be specified. The element g_N will be also specified later. We always assume that

$$\|g_k\| = 1, \quad k = N, \dots. \tag{2.3}$$

We complete the inductive definition of sequences $\{x_n\}_{n=N}^\infty$ and $\{g_n\}_{n=N+1}^\infty$ by setting

$$x_{n+1} := x_n - \langle x_n, g_{n+1}\rangle g_{n+1}. \tag{2.4}$$

We will not specify the sequences $\{\gamma_k\}$, $\{h_k\}$, $\{\xi_k\}$, $0\le\xi_k\le 1$, in this section and get here some general formulas under assumption that the above sequences satisfy some conditions. It will be convenient for us to introduce one more sequence

$$q_{n+1} := \langle x_n, g_{n+1}\rangle, \quad n = N,\dots. \tag{2.5}$$

We assume that $h_{n+1}\in \operatorname{Span}\{e_1,\dots,e_n\}$ for all n. This assumption and (2.4) imply

$$a_{n+1,n+1} = -q_{n+1}\xi_{n+1}.$$

We assume that the sequences $\{\gamma_k\}$ and $\{q_k\}$ satisfy the following condition

$$1-\gamma_{n+1}q_{n+1} = \frac{q_{n+1}}{q_n}, \quad n = N+1,\dots. \tag{2.6}$$

Let us list some identities which will be useful later on. Using (2.2) and (2.5) we get from (2.4) that

$$x_{n+1} = x_n - q_{n+1}g_{n+1} = (1-\gamma_{n+1}q_{n+1})x_n - q_{n+1}h_{n+1} - q_{n+1}\xi_{n+1}e_{n+1} = \tag{2.7}$$

$$\frac{q_{n+1}}{q_n}x_n - q_{n+1}h_{n+1} - q_{n+1}\xi_{n+1}e_{n+1}.$$

The relations (2.1) and (2.2) imply

$$q_{n+1} = \langle x_n, g_{n+1}\rangle = \gamma_{n+1}\|x_n\|^2 + \langle x_n, h_{n+1}\rangle. \tag{2.8}$$

We will need formulas for $\langle x_n, g_k\rangle$. It is clear from (2.4) that

$$\langle x_n, g_n\rangle = 0, \quad n = N+1,\dots. \tag{2.9}$$

I. Case $N\le k<n$**.** By (2.6) and (2.7) we get for $n\ge N+2$

$$\langle x_n, g_k\rangle = \langle \frac{q_n}{q_{n-1}}x_{n-1} - q_nh_n, g_k\rangle = \frac{q_n}{q_{n-1}}\langle x_{n-1}, g_k\rangle - q_n\langle h_n, g_k\rangle. \tag{2.10}$$

Repeating (2.10) and using (2.9) we obtain for $k\ge N+1$

$$\langle x_n, g_k\rangle = \frac{q_n}{q_k}\langle x_k, g_k\rangle - q_n\sum_{l=k+1}^{n}\langle h_l, g_k\rangle = -q_n\sum_{l=k+1}^{n}\langle h_l, g_k\rangle. \tag{2.11}$$

Let us choose g_N now. We specify $h_{N+1}=0$ and take $0<\epsilon<1$. Set

$$g_N := \epsilon x_{N+1}\|x_{N+1}\|^{-2} + \xi g_{N+1} \tag{2.12}$$

with ξ such that $\|g_N\|=1$. Then

$$\langle x_{N+1}, g_N\rangle = \epsilon,$$

and $x_{N+1}\in \operatorname{Span}(g_N, g_{N+1})$. By (2.10) similarly to (2.11) we get for $n\ge N+2$

$$\langle x_n, g_N\rangle = \frac{q_n}{q_{N+1}}\langle x_{N+1}, g_N\rangle - q_n\sum_{l=N+2}^{n}\langle h_l, g_N\rangle. \tag{2.13}$$

II. Case $k > n+1 \geq N+2$. We have by (2.2) for $k \geq n+1$

$$\langle x_n, g_{k+1}\rangle = \langle x_n, \gamma_{k+1}x_k + h_{k+1}\rangle = \gamma_{k+1}\langle x_n, x_k\rangle + \langle x_n, h_{k+1}\rangle. \tag{2.14}$$

Next,

$$\langle x_n, x_k\rangle = \langle x_n, x_{k-1} - q_k g_k\rangle = \langle x_n, x_{k-1}\rangle - q_k\langle x_n, g_k\rangle. \tag{2.15}$$

We also have

$$\langle x_n, g_k\rangle = \langle x_n, \gamma_k x_{k-1} + h_k\rangle = \gamma_k\langle x_n, x_{k-1}\rangle + \langle x_n, h_k\rangle \tag{2.16}$$

and

$$\langle x_n, x_{k-1}\rangle = \gamma_k^{-1}(\langle x_n, g_k\rangle - \langle x_n, h_k\rangle). \tag{2.17}$$

Combining (2.15) and (2.17) we get from (2.14)

$$\langle x_n, g_{k+1}\rangle = (\gamma_k^{-1} - q_k)\gamma_{k+1}\langle x_n, g_k\rangle + \langle x_n, h_{k+1} - \frac{\gamma_{k+1}}{\gamma_k}h_k\rangle. \tag{2.18}$$

Using (2.6) we rewrite (2.18)

$$\langle x_n, g_{k+1}\rangle = \frac{\gamma_{k+1}}{\gamma_k}\frac{q_k}{q_{k-1}}\langle x_n, g_k\rangle + \langle x_n, h_{k+1} - \frac{\gamma_{k+1}}{\gamma_k}h_k\rangle. \tag{2.19}$$

We note that (2.4) implies

$$\|x_n\|^2 = \|x_{n+1}\|^2 + \langle x_n, g_{n+1}\rangle^2$$

and

$$q_{n+1}^2 = \|x_n\|^2 - \|x_{n+1}\|^2, \quad n = N, \dots. \tag{2.20}$$

3. Specifications

Our goal is to prove that the procedure described by (2.4) is a Pure Greedy Algorithm with regard to the dictionary $\mathcal{D} = \{g_m\}_{m=N}^{\infty}$ with approriately chosen N and f. We will choose

$$x_N := -N^{-1/2}\|x_N\|\sum_{i=1}^{N} e_i,$$

with the sequence $\{\|x_n\|\}$ specified below and we define g_N by (2.12). With these two starting elements we use the procedures (2.2) and (2.4) to get the sequences $\{x_n\}_{n=N}^{\infty}$ and $\{g_n\}_{n=N}^{\infty}$. We set $f := x_{N+1}$ and prove that for big enough N (2.4) is a realization of PGA. This means that we should check that for any $n \geq N+1$ and any $m \geq N$, $m \neq n+1$, we have

$$|\langle x_n, g_m\rangle| < q_{n+1}. \tag{3.1}$$

We remind that by the definition of $\{q_k\}$ we have

$$\langle x_n, g_{n+1}\rangle = q_{n+1}.$$

We will prove (3.1), considering separately two cases $m > n+1$ and $m < n+1$.

3.1 Numerical sequences. We define sequences $\{\|x_n\|\}_{n=1}^\infty$, $\{q_n\}_{n=2}^\infty$, $\{\gamma_n\}_{n=3}^\infty$ depending on a parameter $0<\beta<1/4$. In parallel we give in square brackets asymptotic estimates for $\beta=1/6$. This will give a construction for Theorem 1.1. We set

$$(3.2)\quad \|x_1\|=1; \quad \|x_n\|^2=(1-\beta)^2(1-2\beta)^{-1}n^{-1+2\beta}, \quad [\frac{25}{24}n^{-2/3}], \quad n=2,3,\dots.$$

Then by (2.20) we have

$$q_n^2=\|x_{n-1}\|^2-\|x_n\|^2\approx(1-\beta)^2n^{-2+2\beta}, \quad n=2,3,\dots,$$

and

$$(3.3)\qquad q_n\approx(1-\beta)n^{-1+\beta}, \quad [\frac{5}{6}n^{-5/6}].$$

From (2.6) we get

$$(3.4)\qquad \gamma_n=q_n^{-1}-q_{n-1}^{-1}\approx n^{-\beta}, \quad [n^{-1/6}].$$

We will also need the following relations

$$(3.5)\qquad \gamma_{n+1}\|x_n\|^2-q_{n+1}\approx\beta(1-\beta)(1-2\beta)^{-1}n^{-1+\beta}, \quad [\frac{5}{24}n^{-5/6}].$$

$$(3.6)\qquad \frac{\gamma_{n+1}}{\gamma_n}-\frac{q_{n+1}}{q_n}\approx(1-2\beta)n^{-1}, \quad [\frac{2}{3}n^{-1}].$$

$$(3.7)\qquad \frac{\gamma_{n+1}}{\gamma_n}\frac{q_n}{q_{n-1}}=1-(1+o(1))n^{-1}.$$

We will use the following numerical estimate for $\beta=1/6$

$$(3.8)\qquad (1-\beta)(1-2\beta)^{-1/2}=5(24)^{-1/2}\le 1.021.$$

3.2 The sequence $\{h_n\}$. We already agreed to set $h_{N+1}=0$. Let for $n\ge N+2$ define h_n as follows

$$(3.9)\qquad h_n:=\frac{\alpha_n}{n-1}\Phi_n, \quad 0<\alpha_n<\alpha_0<1/2,$$

with

$$\Phi_n:=\sum_k\phi_n(k)e_k$$

where $\phi_n(t):=\chi_{(an,n-1)}(t)$ is the characteristic function of $(an,n-1)$ with a parameter $0<a<1$. In the following numerical estimates we set $a=0.05$. The sequence $\{\alpha_n\}$, $0<\alpha_n<\alpha_0<1/2$, will be chosen later. Then we have

$$(3.10)\qquad \|\Phi_n\|\approx(1-a)^{1/2}n^{1/2}, \quad [0.95^{1/2}n^{1/2}];$$

and

$$\|\Phi_n\|_1 := \sum_k |\phi_n(k)| \approx (1-a)n, \quad [0.95n]. \tag{3.11}$$

It is clear from (2.2), (2.3) and definition of Φ_n that $\xi_n \to 1$ as $n \to \infty$. This and (2.7) imply that for big n we will have

$$a_{n,k} \le -(1+o(1))q_n, \quad n,k \ge N_1, \tag{3.12}$$

and

$$A_m := |\langle x_m, \Phi_{m+1}\rangle| = -\langle x_m, \Phi_{m+1}\rangle \ge (1+o(1))q_m\|\Phi_{m+1}\|_1, \quad [0.79m^{1/6}]. \tag{3.13}$$

The sequence $\{\alpha_n\}$ should satisfy the relation

$$q_{m+1} = \langle x_m, g_{m+1}\rangle = \gamma_{m+1}\|x_m\|^2 + \frac{\alpha_{m+1}}{m}\langle x_m, \Phi_{m+1}\rangle$$

and by (3.5) and (3.13)

$$\frac{\alpha_{m+1}}{m} = A_m^{-1}(\gamma_{m+1}\|x_m\|^2 - q_{m+1}) \le \tag{3.14}$$

$$(1+o(1))((1-a)mq_m)^{-1}\beta(1-\beta)(1-2\beta)^{-1}m^{-1+\beta}, \quad [\frac{5}{19}m^{-1}].$$

Let us specify $\alpha_0 := 0.27$ for $\beta = 1/6$. We will also need an estimate for $A_m - A_{m-1}$. Denote

$$g'_{n+1} := g_{n+1} - \xi_{n+1}e_{n+1} = \gamma_{n+1}x_n + h_{n+1}. \tag{3.15}$$

Taking into account that coordinates of x_n are negative (see (3.12)) and coordinates of h_{n+1} are positive and less in absolute value than the corresponding coordinates of x_n (see (3.9)) we get that the nonzero coordinates $(g'_{n+1})_k$ of g'_{n+1} are negative and

$$\|g'_{n+1}\| \le \gamma_{n+1}\|x_n\|, \quad (g'_{n+1})_k \le -(q_n\gamma_n - \alpha_0/n), \quad k \ge an. \tag{3.16}$$

We have

$$A_{m-1} - A_m = \langle x_m, \Phi_{m+1}\rangle - \langle x_{m-1}, \Phi_m\rangle = \tag{3.17}$$

$$\langle x_{m-1}, \Phi_{m+1} - \Phi_m\rangle - q_m\langle g_m, \Phi_{m+1}\rangle.$$

All coordinates of x_{m-1} are negative. Therefore

$$-(1+o(1))q_m = a_{m-1,m-1} \le \langle x_{m-1}, \Phi_{m+1} - \Phi_m\rangle \le \tag{3.18}$$

$$a_{m-1,m-1} + \|x_m\|'_\infty = -(1+o(1))q_m + \|x_m\|'_\infty,$$

where

$$\|x_m\|'_\infty := \max_{k \ge am} |a_{m,k}|.$$

From (2.7) and definition of $\{h_n\}$ we get

$$\|x_m\|'_\infty \le (1+o(1))q_m(1+\alpha_0 \max_{am\le k\le m} \sum_{l=k}^{k/a} l^{-1}) \le \tag{3.19}$$

$$(1+o(1))q_m(1+\alpha_0 \ln 1/a), \quad [1.51m^{-5/6}].$$

Thus we get from (3.18) for $\beta = 1/6$ and big enough N

$$-0.84m^{-5/6} < \langle x_{m-1}, \Phi_{m+1} - \Phi_m\rangle < 0.68m^{-5/6}. \tag{3.20}$$

We get from (3.16)

$$|\langle g_m, \Phi_{m+1}\rangle| \le \|g'_m\|\|\Phi_{m+1}\| \le \gamma_m \|x_{m-1}\|\|\Phi_{m+1}\| \approx \tag{3.21}$$

$$(1-a)^{1/2}(1-\beta)(1-2\beta)^{-1/2}, \quad [1].$$

On the other hand by (3.16) we get

$$\langle g_m, \Phi_{m+1}\rangle \le -(1-a)m(q_m\gamma_m - \alpha_0/m), \quad [-0.53]. \tag{3.22}$$

Thus from (3.17), (3.20), (3.21), and (3.22) we get for big N

$$-0.4m^{-5/6} \le A_{m-1} - A_m \le 1.52m^{-5/6}. \tag{3.23}$$

Further,

$$A_m^{-1} - A_{m-1}^{-1} = (A_{m-1} - A_m)(A_{m-1}A_m)^{-1}. \tag{3.24}$$

Using (3.13) we get for $\beta = 1/6$ from (3.24) and (3.23) that

$$-0.641m^{-7/6} < A_m^{-1} - A_{m-1}^{-1} < 2.43m^{-7/6}. \tag{3.25}$$

We proceed now to estimate $|\langle x_n, g_{m+1}\rangle|$. Consider first the case $m > n$. We will prove (3.1) in this case by induction using the following representation formula (2.19)

$$\langle x_n, g_{m+1}\rangle = \frac{\gamma_{m+1}}{\gamma_m}\frac{q_m}{q_{m-1}}\langle x_n, g_m\rangle + \langle x_n, h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m}h_m\rangle. \tag{3.26}$$

By (3.7) and induction assumption we get for the first term

$$|\frac{\gamma_{m+1}}{\gamma_m}\frac{q_m}{q_{m-1}}\langle x_n, g_m\rangle| \le (1-(1+o(1))m^{-1})q_{n+1}. \tag{3.27}$$

Thus we need to prove that there exists $\delta > 0$ such that

$$|\langle x_n, h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m}h_m\rangle| \le (1-\delta)m^{-1}q_{n+1}. \tag{3.28}$$

We have

$$\langle x_n, h_{m+1} - \frac{\gamma_{m+1}}{\gamma_m} h_m\rangle = \tag{3.29}$$

$$\langle x_n, \frac{\alpha_{m+1}}{m}(\Phi_{m+1} - \Phi_m)\rangle + \langle x_n, \Phi_m(\frac{\alpha_{m+1}}{m} - \frac{\gamma_{m+1}}{\gamma_m}\frac{\alpha_m}{m-1})\rangle =: a_1 + a_2.$$

We have by (3.19)

$$|a_1| = |\langle x_n, \frac{\alpha_{m+1}}{m}(\Phi_{m+1} - \Phi_m)\rangle \leq \tag{3.30}$$

$$\|x_n\|'_\infty \frac{\alpha_{m+1}}{m} \leq (1 + o(1))(1 + \alpha_0 \ln 1/a) q_{n+1} \alpha_0 m^{-1}, \quad [0.41 m^{-1} n^{-5/6}].$$

For a_2 we have

$$|a_2| \leq \|x_n\| \|\Phi_m\| |\frac{\alpha_{m+1}}{m} - \frac{\gamma_{m+1}}{\gamma_m}\frac{\alpha_m}{m-1}|. \tag{3.31}$$

Using (3.14) we get

$$\frac{\alpha_{m+1}}{m} - \frac{\gamma_{m+1}}{\gamma_m}\frac{\alpha_m}{m-1} = (A_m^{-1} - A_{m-1}^{-1})(\gamma_{m+1}\|x_m\|^2 - q_{m+1}) + \tag{3.32}$$

$$A_{m-1}^{-1}(\gamma_{m+1}(\|x_m\|^2 - \|x_{m-1}\|^2) - q_{m+1} + \frac{\gamma_{m+1}}{\gamma_m} q_m) =: \sigma_1 + \sigma_2.$$

Taking into account (3.5) we get from (3.25) the following estimates for σ_1

$$-0.14 m^{-2} \leq \sigma_1 \leq 0.51 m^{-2}. \tag{3.33}$$

Let us proceed to σ_2 now. We have

$$\gamma_{m+1}(\|x_m\|^2 - \|x_{m-1}\|^2) - q_{m+1} + \frac{\gamma_{m+1}}{\gamma_m} q_m \approx -\beta(1-\beta) m^{-1}. \tag{3.34}$$

Using this, (3.13) and the following inequality for A_m

$$A_m \leq \|x_m\| \|\Phi_{m+1}\| \approx (1-\beta)(1-2\beta)^{-1/2}(1-a)^{1/2} m^\beta, \quad [m^{1/6}],$$

we get

$$-0.18 m^{-2} m^{-2} \leq \sigma_2 \leq -0.13 m^{-2}. \tag{3.35}$$

Thus

$$-0.32 m^{-2} \leq \sigma_1 + \sigma_2 \leq 0.38 m^{-2}.$$

and

$$|a_2| \leq |\sigma_1 + \sigma_2| \|x_n\| \|\Phi_m\| \leq 0.38 m^{-3/2} n^{-1/3}. \tag{3.36}$$

Combining (3.30) with (3.36) we get

$$|a_1| + |a_2| \leq 0.79 m^{-1} n^{-5/6} < 0.99 m^{-1} q_{n+1}. \tag{3.37}$$

This inequality holds for all $m > n$ and big enough N. Using (3.7) we inductively get from (3.26) and (3.37) that for all $m > n$

$$|\langle x_n, g_{m+1}\rangle| < q_{n+1}. \tag{3.38}$$

We proceed now to the case $m < n$. By (2.11) and (2.13) we need to estimate

$$|\sum_{l=k+1}^{n} \langle h_l, g_k\rangle|, \quad k = N+1, \dots.$$

We use the definition $h_l = \frac{\alpha_l}{l-1}\Phi_l$ and the inequality $0 < \alpha_l \le \alpha_0$ which holds if N is big enough. If $l > 20k$ then $\langle h_l, g_k\rangle = 0$. For $l = k+1$ we have

$$|\langle h_l, g_k\rangle| = |\langle h_l, g'_k\rangle| \le \|h_l\|\|g'_k\| \le Ck^{-1}.$$

For $k+2 \le l \le 20k$ we have

$$\langle h_l, g_k\rangle = \frac{\alpha_l}{l-1}(\xi_k - |\langle \Phi_l, g'_k\rangle|) \tag{3.39}$$

and by (3.16)

$$|\langle \Phi_l, g'_k\rangle| \le k^{1/2}\|g'_k\| \le 1.5.$$

Thus

$$|\langle h_l, g_k\rangle| < 0.27 l^{-1}, \quad l \ge k+2,$$

and therefore

$$\sum_{l=k+1}^{20k} |\langle h_l, g_k\rangle| \le (1+o(1))0.27\ln 20 < 0.81. \tag{3.40}$$

Thus we have for $m \le n$

$$|\langle x_n, g_m\rangle| \le (0.81 + 0.05)q_{n+1} < q_{n+1} \tag{3.41}$$

for $\epsilon = 0.05q_{N+1}$ and big N.

Let us complete the construction of a counterexample. Let N be big enough to ensure (3.38) and (3.41) for $n, m \ge N$. Then we set

$$x_N := -N^{-1/2}5(24)^{-1/2}N^{-1/3}\sum_{i=1}^{N} e_i$$

and

$$g_{N+1} := q_{N+1}x_N\|x_N\|^{-2} + \xi_{N+1}e_{N+1}$$

with q_{N+1} chosen from (3.3) with $\beta = 1/6$ and $\xi_{N+1} > 0$ such that $\|g_{N+1}\| = 1$. We define now $\{x_n\}_{n\ge N+1}$ and $\{g_n\}_{n\ge N+2}$ by (2.2) and (2.4) with $\{\gamma_n\}$, $\{h_n\}$ specified in Section 3 with $\beta = 1/6$. We set $\epsilon = 0.05q_{N+1}$ and define

$$g_N := \epsilon x_{N+1}\|x_{N+1}\|^{-2} + \xi g_{N+1}, \quad \|g_N\| = 1.$$

Consider now $f := x_{N+1}$ and $\mathcal{D} := \{g_N, g_{N+1}, \dots\}$. We have that $f \in \mathrm{Span}(g_N, g_{N+1})$ and therefore $|f|_{\mathcal{A}_1(\mathcal{D})} \le C(N)$. By (3.38) and (3.41) the PGA with regard to $\mathcal{D}$ applied to f will exactly realize the iterative process (2.4). Thus

$$f_m = x_{N+1+m}$$

and

$$\|f_m\|^2 = \frac{25}{24}(N+1+m)^{-2/3}.$$

This implies

$$\|f_m\| \ge C_1(N)m^{-1/3}$$

which completes the proof of Theorem 1.1.

4. Proof of Theorem 1.3

The construction of $\mathcal{D}_t$ is the same as in Section 2 of [LT]. It uses the Equalizer procedure. Let H be a Hilbert space with an orthonormal basis $\{e_j\}_{j=1}^{\infty}$. For two elements e_i, e_j, $i \neq j$, and for a positive number $t \leq 1/3$ we define the procedure which we call "equalizer" and denote $E(e_i, e_j, t)$.

Equalizer $E(e_i, e_j, t)$**.** Denote $f_0 := e_i$ and $g_1 := \alpha_1 e_i - (1-\alpha_1^2)^{1/2} e_j$ with $\alpha_1 := t$. Then $\|g_1\| = 1$ and $\langle f_0, g_1 \rangle = t$. We define the sequences $f_1, \dots, f_N$; $g_2, \dots, g_N$; $\alpha_2, \dots, \alpha_N$ inductively:

$$f_n := f_{n-1} - \langle f_{n-1}, g_n \rangle g_n; \quad g_{n+1} := \alpha_{n+1} e_i - (1-\alpha_{n+1}^2)^{1/2} e_j$$

with α_{n+1} satisfying

$$\langle f_n, g_{n+1} \rangle = t, \quad n = 1, 2, \dots .$$

Let $f_n = a_n e_i + b_n e_j$ and $N := N_t$ be the number such that

$$a_{N-1} - b_{N-1} \geq \sqrt{2}t, \quad a_N - b_N < \sqrt{2}t.$$

Then we modify the N-th step as follows. We take $g_N := 2^{-1/2}(e_i - e_j)$ and

$$f_N = f_{N-1} - \langle f_{N-1}, g_N \rangle g_N.$$

It is clear that then $a_N = b_N$ and

$$t \leq \langle f_{N-1}, g_N \rangle \leq 2t.$$

We list here the following simple relations

$$a_{n+1} = a_n - t\alpha_{n+1}; \quad b_{n+1} = b_n + t(1-\alpha_{n+1}^2)^{1/2}, \quad n < N-1;$$

$$a_{n+1} - b_{n+1} = a_n - b_n - t(\alpha_{n+1} + (1-\alpha_{n+1}^2)^{1/2}), \quad n < N-1; \tag{4.1}$$

$$\|f_{n+1}\|^2 = \|f_n\|^2 - t^2, \quad n < N-1.$$

Relation (4.1) and the inequality $1 \leq x + (1-x^2)^{1/2} \leq 2^{1/2}$, $0 \leq x \leq 1$, imply that

$$N \leq 1/t \tag{4.2}$$

and

$$\|f_N\|^2 \geq \|f_{N-1}\|^2 - 4t^2 \geq \|f\|^2 - t - 3t^2.$$

It is clear that $E(e_i, e_j, t)$ is a WGA with regard to the dictionary $e_i, g_1, g_2, \dots, g_N$ with the "weakness" parameter t.

We define WGA and a dictionary $\mathcal{D}_t$ as follows. We begin with $f := e_1$ and apply $E(e_1, e_2, t)$. After $N_t \geq 1$ steps we get $g_1^0, \dots, g_{N_t}^0$ and

$$f^1 = c_1(e_1 + e_2)$$

with the property

$$\|f^1\|^2 \geq \|f\|^2 - (t+3t^2) = \|f\|^2(1-t-3t^2).$$

We use now $E(e_1,e_3,t)$ and $E(e_2,e_4,t)$. After $2N_t \geq 2$ steps we obtain $g_1^1,\dots,g_{2N_t}^1$ and

$$f^2 = c_2(e_1+\dots+e_4)$$

with the property

$$\|f^2\|^2 \geq \|f^1\|^2(1-t-3t^2) \geq \|f\|^2(1-t-3t^2)^2.$$

After s iterations we get

$$f^s = c_s(e_1+\dots+e_{2^s})$$

and apply $E(e_i,e_{i+2^s},t)$, $i=1,2,\dots,2^s$. We make $2^sN_t \geq 2^s$ steps and get $g_1^s,\dots,g_{2^sN_t}^s$ and

$$f^{s+1} = c_{s+1}(e_1+\dots+e_{2^{s+1}})$$

with the property

$$\|f^{s+1}\|^2 \geq \|f\|^2(1-t-3t^2)^{s+1}.$$

Thus we have

$$\|f - G_{2^s}^t(f,\mathcal{D}_t)\| \geq (1-t-3t^2)^s, \quad s=1,2,\dots, \tag{4.3}$$

with the dictionary

$$\mathcal{D}_t = \bigcup_{k\in\mathbb{N}} e_k \cup \bigcup_{s\geq 0; 1\leq l\leq 2^sN_t} g_l^s.$$

The relation (4.3) and monotonicity of $\|f - G_m^t(f,\mathcal{D}_t)\|$ prove the Theorem 1.3.

Remark 4.1. *The estimate (1.5) implies that for small t the parameter a in (1.6) can be taken close to $1/4$. The inequality (4.3) implies that the parameter b in (1.7) can be taken close to $(\ln 2)^{-1}$.*

References

[DT] R.A. DeVore and V.N. Temlyakov, *Some remarks on Greedy Algorithms*, Advances in Computational Mathematics **5** (1996), 173–187.

[J] L. Jones, *On a conjecture of Huber concerning the convergence of projection pursuit regression*, The Annals of Statistics **15** (1987), 880–882.

[KT] S.V. Konyagin and V.N. Temlyakov, *Rate of convergence of Pure Greedy Algorithm*, East J. on Approx. **5** (1999), 493–499.

[L] E.D. Livshitz, *On the rate of convergence of greedy algorithm*, Manuscript (2000).

[LT] E.D. Livshitz and V.N. Temlyakov, *On convergence of Weak Greedy Algorithms,*, IMI-Preprint **13** (2000), 1–9.

[T] V.N. Temlyakov, *Weak Greedy Algorithms*, Advances in Computational Mathematics **12** (2000), 213–227.